Express Mailing# EV221423084US

KNOWLEDGE-BASED FLEXIBLE NATURAL SPEECH DIALOGUE SYSTEM

[0001] This application claims priority to U.S. Provisional Application Serial No. 60/432,569, filed December 11, 2002

BACKGROUND OF THE INVENTION

[0002] The present invention is mainly directed to a knowledge support and flexible dialogue control system.

[0003] Automatic telephone conversation systems, which are activated in response to a user's request through speech for providing information and service, are well known in the IT industry. An automatic telephone conversation system may contain the components such as a speech recognition engine, a text to speech engine, a natural language understanding engine, a dialogue control engine and some business servers. The dialogue control system may further include a dialogue grammar engine for modeling dialogue structures and for guiding the procedure of satisfying user needs.

[0004] Several known telephone conversation systems include a dialogue control and dialogue grammar system. The dialogue control system could consist of user intention determination based on dialogue act sequencing. A controller, which is connected to one or a combination of these dialogue grammar models, controls the system dialogue moves in accordance with the user intention decided at a point of the dialogue. In response to the understood user intention, one or more deployment aspects of the telephone conversation system, such as a database server, may be accessed. A

conversation system with flexible aspects of dialogue moves control is commonly referred to as a "mixed-initiative" dialogue system.

[0005] Dialogue grammar and dialogue control engines are key components of mixed-initiative telephone conversation systems. There are several types such systems but many of them suffer from serious shortcomings. A system that relies on a generative dialogue act grammar may hardly capture the full flexibility of the conversation flow, for instance. A system that retains the interactive information between the user and the system in the local grammar tree recently generated suffers from the inflexibility of knowledge representation as well as limitation of the locality of the temporal scope. A system that relies solely on the grammar structure to capture the user's knowledge, intention or indication cannot account for other aspects of the knowledge structure, such as the ontological structure, for instance.

SUMMARY OF THE INVENTION

[0006] In an automatic conversation system according to the present invention, flexibilities of the conversation structure, inherent in mixed-initiative mode for dealing with complex user request, are well-managed because the knowledge structures involved are represented by additional, powerful knowledge representation tools, and because the context information is retained by more specific data structures, which covers larger temporal scopes by the logic of the conversation, rather than by a fixed locality of the grammar flow. This invention provides a simple yet reliable method to compensate for these factors to enable more powerful conversation engines with mixed-initiative capabilities.

[0007] The present invention is directed to a novel knowledge-based natural speech dialogue system. In accordance with the present invention, a knowledge-based natural speech dialogue system provides: (i) a knowledge support system, (ii) a flexible dialogue management system, and (iii) a context information system.

[0008] In accordance with a preferred embodiment of the present invention, the knowledge support module comprises: (a) a knowledge representation database, which supports the knowledge in the form of an ontology and features of entities and activities, (b) an interface to the knowledge database, which accesses the knowledge database and gets relevant information based on user requests.

[0009] As for the flexible dialogue management module, it comprises: (a) an interface to the speech recognition engine, through which the recognized words of the user's speech are obtained and further processed, (b) an interface to the natural language understanding engine, to which the recognized words are sent for semantic processing and from which the conceptual meanings of the utterances are obtained, (c) an interface to the knowledge support module in order to obtain needed information, (d) an interface to the context information module in order to obtain information of previous sentences in the dialogue and to store necessary information of the current sentence for use by later stages, and (d) a rule engine in which to store dialogue act strategies which controls the normal flow of conversation according to general principles of verbal interactions.

[0010] The context information module comprises: (a) a data structure that is used to store structured information of some foregoing interactions, and (b) a set of updating instructions, which is used by the dialogue management module for accessing and storing information in the context information data structure.

[0011] The present invention has no restrictions on the type of knowledge database to be used. Any type of database can be used as long as it provides with the system with the functionality it is supposed to provide it with.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Other advantages of the present invention can be understood by reference to the following detailed description when considered in connection with the accompanying drawings wherein:

[0013] FIG. 1 is a schematic block diagram of the flexible natural speech dialogue system (FNDS).

[0014] FIG. 2 is a flow chart of the knowledge support algorithm.

[0015] FIG. 3 is a flow chart of the dialogue management algorithm.

[0016] FIG. 4 is a flow chart of the context information update algorithm.

[0017] FIG. 5 is a schematic of a computer on which the flexible natural speech dialogue system can be implemented.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0018] Referring to FIG. 1, in the flexible natural speech dialogue system (FNDS), the conversation control system is the core of the FNDS and communicates with other servers, such as text-to-speech 410, speech recognition 412, telephone interface 414, natural language understanding 416, business servers 418. The core dialogue management system comprises knowledge representation database 422, knowledge base interface 424, dialogue act logic unit 426, context information storage 420 and context

information interface 428. The flexible dialogue control core system receives recognition results, calls natural language understanding unit to obtain the conceptual representation. Based on the conceptual representation the control unit calls context information for further interpretation of the meaning. Then the control unit calls knowledge support unit 422, 424 and dialogue act rules 426 in order to decide the response to the user. In case clarification or repair is needed, it initiates a sub-dialogue based on dialogue act principles; The core control unit then generates responses to the user by calling the TTS engine. In case some other services are requested, such as search or update databases, it will access the business databases as well.

[0019] FIG. 2 provides a flow chart of the knowledge support algorithm. Request for knowledge base search 512 comes from the dialogue act control unit. (ref. Fig. 1) The judgment unit 514 decides whether it is a request for objects and their properties 516 or for processes and their relations. 518. At decision point 520, if the property is found, results will go out at return 524, otherwise, the parent concept will be searched for the property. At decision point 526, if the relation information is found, it will be sent out. Otherwise, using any nearest neighbor search algorithm for similar concepts, the search is re-directed to this concept. Both of the re-direction procedures are iterative;

[0020] FIG. 3 provides a flow chart of the dialogue management algorithm. This unit controls the information flow of the conversation system. Recognized words 622 from the speech recognition engine are sent to natural language understanding engine at procedure 624. The result of conceptual understanding 625 is sent to context rule engine for further interpretation, such as the hidden implicature of the utterance by

procedure 630. Once the interpretation is obtained, the knowledge support engine is called at procedure 632 to search relevant knowledge as the basis for generating responses. At decision point 634 TTS engine may be called to generate speech response to the user. At decision point 638 business servers may be called to perform some requested actions for the user, before control is transferred to the next dialogue turn;

[0021] FIG. 4 provides a flow chart of the context information update algorithm. The natural language understanding result 720 is examined at decision point 730 with respect to context information structure (ref. 420 in Fig.1). At the decision point 750 it is examined whether enough information is contained in the concept structure. If enough information is found, the context information unit generates a normal output 770; otherwise it sets a check for clarification with the user. If the previous context is in checked state, it is examined whether this check is a yes/no question or not 740. With the yes/no check, if the expected answer is obtained, a normal output is generated 782. Otherwise a check is set up again. In case of other checks, again a decision is made at 780 to judge whether expected answer is obtained or not;

[0022] FIG. 5 is a schematic for a computer 10 on which the fuzzy natural language concept system described above can be implemented. The computer 10 includes a CPU 12, memory 14, such as RAM, and storage 16, such as a hard drive, RAM, ROM or any other optical, magnetic or electronic storage. The computer 10 further includes an input 18 for receiving the speech input, such as over a telephone line, and an output 20 for producing the responsive speech output, such as over the telephone line. The computer 10 may also include a display 22. The algorithms, software and databases described above with respect to Figs. 1-4 are implemented on the computer 10

and are stored in the memory 14 and/or storage 16. The computer 10 is suitably programmed to perform the steps and algorithms described herein.

[0023] From the above description of a preferred embodiment of the invention, those skilled in the art will perceive improvements, changes and modifications. Such improvements, changes and modifications within the skill of the art are intended to be covered by the appended claims.